ELSEVIER
Statistics & Probability Letters 52 (2001) 421–426
STATISTICS & PROBABILITY LETTERS
www.elsevier.nl/locate/stapro

On the Fisher information in randomly censored data

Gang Zheng[a,*], Joseph L. Gastwirth[b]

[a] *Office of Biostatistics Research, National Heart, Lung, and Blood Institute, Room 8223, Rockledge 2, 6701 Rockledge Drive, Bethesda, MD 20892-7938, USA*

[b] *Department of Statistics, George Washington University, Washington, DC 20052, USA*

Received August 2000; received in revised form January 2001

Abstract

After expressing the Fisher information in randomly censored data in terms of hazard rates, we examine the loss of Fisher information under the Koziol–Green random censoring model. Because the same parameter enters both the survival and censoring distributions in this model, the censored data may contain more Fisher information than the uncensored data.

Keywords: Hazard rate; Koziol–Green model; Loss of Fisher information

1. Introduction

Let X and Y be independent random variables with distribution functions $F(x;\theta)$ and $G(x;\theta)$, respectively, where $x \in E = (a,b)$, a finite or infinite interval, and $\theta \in \Theta \subset R^1$. Under right random censorship, we only observe (Z,δ), where $Z = \min(X,Y)$ and $\delta = I(X \leqslant Y)$ and $I(A)$ is an indicator function. The full likelihood for (Z,δ) is

$$L(z,\delta) = \{f(z;\theta)\bar{G}(z;\theta)\}^{\delta}\{g(z;\theta)\bar{F}(z;\theta)\}^{1-\delta},$$

where $\bar{F} = 1 - F$ and $\bar{G} = 1 - G$. The Fisher information about θ contained in (Z,δ) is defined by $I^{Z,\delta}(\theta) = E\{\partial/\partial\theta \log L(Z,\delta)\}^2$, which can be expressed (Miller, 1983; Prakasa Rao, 1995) as

$$I^{Z,\delta}(\theta) = \int_E \left\{\frac{\partial}{\partial\theta}\log(f\bar{G})\right\}^2 f\bar{G}\,\mathrm{d}x + \int_E \left\{\frac{\partial}{\partial\theta}\log(g\bar{F})\right\}^2 g\bar{F}\,\mathrm{d}x. \tag{1}$$

* Corresponding author.
E-mail addresses: zhengg@nhlbi.nih.gov (G. Zheng), jlgast@research.circ.gwu.edu (J.L. Gastwirth).

PII: S0167-7152(01)00025-6

Denote the hazard rate of a random variable X by $\lambda_X = f/\bar{F}$. Under suitable regularity conditions, Efron and Johnstone (1990) studied the following identity about the Fisher information in a random variable X,

$$I^X(\theta) = \int_E \left(\frac{\partial}{\partial\theta}\log f\right)^2 f\,\mathrm{d}x = \int_E \left(\frac{\partial}{\partial\theta}\log \lambda_X\right)^2 f\,\mathrm{d}x. \tag{2}$$

Using an analogous expression of $I^{Z,\delta}(\theta)$ in terms of the hazard rates of X and Y, we examine the loss of Fisher information due to random censoring.

2. Fisher information in randomly censored data

In addition to the standard regularity conditions under which the Fisher information for F and G exists, we also assume that $(\partial/\partial\theta \log \bar{F})^2\bar{F} \to 0$, $(\partial/\partial\theta \log \bar{G})^2\bar{G} \to 0$ as $x \to b$.

Theorem 1.

$$I^{Z,\delta}(\theta) = I^X(\theta) - \int_E \left(\frac{\partial}{\partial\theta}\log\lambda_X\right)^2 fG\,\mathrm{d}x + I^Y(\theta) - \int_E \left(\frac{\partial}{\partial\theta}\log\lambda_Y\right)^2 gF\,\mathrm{d}x. \tag{3}$$

Proof. Denote $\partial F/\partial\theta$ by F'_θ and $\partial f/\partial\theta$ by f'_θ. From the first term on the right of (1),

$$\int_E \left\{\frac{\partial}{\partial\theta}\log(f\bar{G})\right\}^2 f\bar{G}\,\mathrm{d}x$$

$$= I^X(\theta) - \int_E \left(\frac{\partial}{\partial\theta}\log f\right)^2 fG\,\mathrm{d}x - 2\int_E \bar{G}'_\theta\,\mathrm{d}\bar{F}'_\theta + \int_E \left(\frac{\partial}{\partial\theta}\log\bar{G}\right)^2 \bar{G}\,\mathrm{d}F. \tag{4}$$

Integrating by parts, the last two terms on the right of (4) become

$$-2\int_E g'_\theta \bar{F}'_\theta\,\mathrm{d}x - \int_E \left\{-\frac{2\bar{G}'_\theta g'_\theta}{\bar{G}} + \frac{(\bar{G}'_\theta)^2 g}{\bar{G}^2}\right\} F\,\mathrm{d}x.$$

Similarly, we obtain

$$\int_E \left\{\frac{\partial}{\partial\theta}\log(g\bar{F})\right\}^2 g\bar{F}\,\mathrm{d}x = I^Y(\theta) - \int_E \left(\frac{\partial}{\partial\theta}\log g\right)^2 gF\,\mathrm{d}x + 2\int_E g'_\theta\bar{F}'_\theta\,\mathrm{d}x$$

$$- \int_E \left\{-\frac{2\bar{F}'_\theta f'_\theta}{\bar{F}} + \frac{(\bar{F}'_\theta)^2 f}{\bar{F}^2}\right\} G\,\mathrm{d}x.$$

Thus substituting the above result and (4) into (1), we can obtain (2). □

From (2) and (3), we have,

$$I^X(\theta) = \int_E \left(\frac{\partial}{\partial\theta}\log\lambda_X\right)^2 f\,\bar{G}\,\mathrm{d}x + \int_E \left(\frac{\partial}{\partial\theta}\log\lambda_X\right)^2 f\,G\,\mathrm{d}x, \tag{5}$$

$$I^{Z,\delta}(\theta) = \int_E \left(\frac{\partial}{\partial\theta}\log\lambda_X\right)^2 f\,\bar{G}\,\mathrm{d}x + \int_E \left(\frac{\partial}{\partial\theta}\log\lambda_Y\right)^2 g\,\bar{F}\,\mathrm{d}x. \tag{6}$$

Comparing (5) and (6), we find that the first term on the right side of (5) is also in (6) but the second term, $\int_E(\partial/\partial\theta \log\lambda_X)^2 f\,G\,\mathrm{d}x$, is not. Instead, $I^{Z,\delta}(\theta)$ contains a new term, $\int_E(\partial/\partial\theta\log\lambda_Y)^2 g\bar{F}\,\mathrm{d}x$, which vanishes when G is independent of θ. Zheng (2001) studied the factorization of hazard function, $\lambda_X = u(x)v(\theta)$, as is the case when X is Weibull. In this case,

$$\int_E\left(\frac{\partial}{\partial\theta}\log\lambda_X\right)^2 f\,\bar{G}\,\mathrm{d}x = I^X(\theta)\,\mathrm{pr}(X<Y).$$

This implies that the portion of the total Fisher information contained in the censored data is just the proportion of the observations that are not censored. This motivates the following definitions:

Under right random censorship, the retained $(I_{\mathrm{R}}^X(\theta))$ and lost $(I_{\mathrm{L}}^X(\theta))$ Fisher information about θ in X due to censoring by the variable Y are defined, respectively, as

$$I_{\mathrm{R}}^X(\theta) = \int_E\left(\frac{\partial}{\partial\theta}\log\lambda_X\right)^2 f\,\bar{G}\,\mathrm{d}x \quad\text{and}\quad I_{\mathrm{L}}^X(\theta) = \int_E\left(\frac{\partial}{\partial\theta}\log\lambda_X\right)^2 f\,G\,\mathrm{d}x.$$

Using these definitions, $I^X(\theta) = I_{\mathrm{R}}^X(\theta) + I_{\mathrm{L}}^X(\theta)$ and $I^{Z,\delta}(\theta) = I_{\mathrm{R}}^X(\theta) + I_{\mathrm{R}}^Y(\theta)$.

A property of Fisher information in randomly censored data is given below.

Theorem 2. *$I^{Z,\delta}(\theta) \geqslant I^Z(\theta)$ where the equality holds if and only if λ_Y/λ_X is independent of θ.*

Proof. Using (1),

$$I^{Z,\delta}(\theta) - I^Z(\theta) = \int_E \frac{(f\bar{G})(g\bar{F})}{f\bar{G}+g\bar{F}}\left\{\frac{\partial}{\partial\theta}\log\left(\frac{\lambda_X}{\lambda_Y}\right)\right\}^2 \mathrm{d}x \geqslant 0,$$

where the equality holds if and only if λ_Y/λ_X is independent of θ. □

3. Loss of Fisher information in the Koziol–Green model

If the censoring distribution G does not involve the parameter θ, then $I^{Z,\delta}(\theta) = I_{\mathrm{R}}^X(\theta)$ and the loss of Fisher information is $I_{\mathrm{L}}^X(\theta)$. In some applications, however, the parameter of interest also affects G. This occurs in the Koziol–Green model of random censoring (KGM), where it is assumed that $\bar{G} = \bar{F}^\beta$, i.e., $\lambda_Y = \beta\lambda_X$, where $\beta > 0$ is a constant (e.g., Koziol and Green, 1976; Csörgő and Horváth, 1981; Chen et al., 1982; Cheng and Lin, 1987; Stute, 1992; Pawlitschko, 1999). Let $I^Z(\theta)$ be Fisher information in $Z = \min(X,Y)$. As the density of Z is $h = f\bar{G} + g\bar{F}$, $I^Z(\theta) = \int_E(\partial/\partial\theta\log h)^2 h\,\mathrm{d}x$. From Theorem 2, under the KGM, $I^{Z,\delta}(\theta) = I^Z(\theta)$.

In their analysis of the efficiency of the nonparametric maximum likelihood estimator (MLE) of the survival function $(\hat{S}(t))$ under the KGM, Hollander et al. (1985) observed that the asymptotic variance of $\hat{S}(t)$ can decrease as the degree of censoring, reflected by the constant β, increases. When G is a function of θ, $I_{\mathrm{R}}^Y(\theta)$ is not 0. In this case, (5) and (6) indicate that the loss of Fisher information should be $I_{\mathrm{L}}^X(\theta) - I_{\mathrm{R}}^Y(\theta)$. Under the KGM, from (5) and (6),

$$I_{\mathrm{L}}^X(\theta) - I_{\mathrm{R}}^Y(\theta) = I^X(\theta) - I^{Z,\delta}(\theta) = (1+\beta)\int_E\left(\frac{\partial}{\partial\theta}\log\lambda_X\right)^2\left(\frac{1}{1+\beta} - \bar{G}\right) f\,\mathrm{d}x.$$

Thus (Z,δ) contains more Fisher information than X if and only if

$$I_{\mathrm{R}}^X(\theta) > \{1/(1+\beta)\}I^X(\theta). \tag{7}$$

Under the KGM, $1/(1+\beta)$ is the probability that an observation is not censored. Abdushukurov and Kim (1987) noticed that, if the KGM holds and $\lambda_X = u(x)v(\theta)$, then no Fisher information is lost in (Z,δ). In this case, (7) is an equality.

Table 1
Fisher information in (Z,δ) under the KGM.[a]

Distribution F	θ	$\beta=0.0$	$\beta=0.5$	$\beta=1.0$	$\beta=2.0$	$\beta=3.0$	$\beta=5.0$	$\beta=20$
Normal	Location	1.000	1.264	1.481	1.827	2.103	2.533	4.140
Logistic	Location	0.333	0.429	0.500	0.600	0.667	0.750	0.913
Gamma	Scale[b]	2.000	2.160	2.277	2.444	2.563	2.725	3.172
Half logistic	Scale	1.430	1.426	1.404	1.351	1.304	1.237	1.086
Weibull	Scale[c]	c^2	c^2	c^2	c^2	c^2	c^2	c^2

[a]Note. $\beta=0$ corresponds to no censoring.
[b]The shape parameter is 2.
[c]The shape parameter is c.

Denote

$$h(x;\theta)=\frac{\partial}{\partial x}\left(\frac{\partial}{\partial\theta}\log\lambda_X\right)^2.$$

Under the KGM,

$$I^{Z,\delta}(\theta)=(1+\beta)\int_E\left(\frac{\partial}{\partial\theta}\log\lambda_X\right)^2 f\bar{G}\,\mathrm{d}x=\left(\frac{\partial}{\partial\theta}\log\lambda_X\right)^2|_{X=a}+\int_E\bar{F}^{1+\beta}h(x;\theta)\,\mathrm{d}x$$

$$=I^X(\theta)-\int_E\bar{F}(1-\bar{F}^\beta)h(x;\theta)\,\mathrm{d}x.$$

This implies the following result.

Theorem 3. *Under the* KGM, *if* $h(x;\theta)\leqslant(\geqslant)0$ *for* $x\in E$ *and* $\theta\in\Theta$, *then* $I^{Z,\delta}(\theta)$ *is a non-decreasing* (*non-increasing*) *function of* β.

Routine calculations show that $h(x;\theta)<0$ for the location parameters of data from either normal or logistic distributions and for the scale parameters of a Gamma distribution with the shape parameter 2. Thus, in these cases, $I^{Z,\delta}(\theta)$ is a strictly increasing function of β. For the scale parameter of the Weibull distribution, with any shape parameter, $h(x;\theta)=0$, hence $I^{Z,\delta}(\theta)=I^X(\theta)=I^Z(\theta)$ for $\beta>0$. On the other hand, for the half logistic distribution with the density $f(x;\theta)=2\exp(-x/\theta)\{1+\exp(-x/\theta)\}^{-2}/\theta$ for $x>0$ and $\theta>0$ (Balakrishnan, 1985), $h(x;\theta)$ changes sign with x. To illustrate the extra information available in the KGM setting, values of the Fisher information in censored data, $I^{Z,\delta}(\theta)$, for the distributions discussed are reported in Table 1. In Table 1, under the KGM, when $\beta=0$, the corresponding $I^{Z,\delta}(\theta)$ is the Fisher information about θ contained in X. When $I^{Z,\delta}(\theta)$ with $\beta=0$ is less than $I^{Z,\delta}(\theta)$ with $\beta>0$, (Z,δ) contains more Fisher information about θ than X. This means that observations from the censoring distribution provide more information about θ than observations from the survival distribution, thus we gain Fisher information due to random censoring under the KGM. Otherwise, the loss of Fisher information occurs due to random censoring even under KGM.

Suppose we estimate $\phi(\theta)$, a function of the parameter θ in $F(x;\theta)$, by an unbiased estimate $\hat{\phi}(Z,\delta)$. Abdushukurov and Kim (1987) showed that, under certain regularity conditions,

$$\mathrm{Var}[\hat{\phi}(Z,\delta)]\geqslant\frac{[\phi'_\theta(\theta)]^2}{I^{Z,\delta}(\theta)}.$$

This implies that, under the KGM, the lower bound of the variance of an unbiased estimate using (Z,δ) would be smaller than that based on X if extra information is available in (Z,δ). If we are interested in estimating the survival function $\bar{F}(x;\theta)$ when both $F(x;\theta)$ and $G(x;\theta)$ are known, from Miller (1983), under regularity conditions, we have

$$\mathrm{Var}_{\mathrm{asy}}[\bar{F}(x;\hat{\theta})] = \frac{[\bar{F}'_{\theta}(x;\theta)]^2}{nI^{Z,\delta}(\theta)},$$

where $\hat{\theta}$ is the MLE based on $(Z_i,\ \delta_i)$, $i = 1,\ \ldots,\ n$. This also implies that the efficiency of the parametric estimate of the survival function can be improved under the KGM for some underlying distribution families. This result is consistent with those of Hollander et al. (1985) and Cheng and Lin (1987), since under some KGM, there is more information about θ in the censoring distribution.

4. Related results

In this paper, we focus on the Fisher information about a single parameter in right censored data. In the multiparameter case $\theta = (\theta_1, \ldots, \theta_m)$, (5) becomes

$$I^{Z,\delta}(\theta_i,\theta_j) = \int_E \left(\frac{\partial}{\partial\theta_i}\log\lambda_X\right)\left(\frac{\partial}{\partial\theta_j}\log\lambda_X\right) f\bar{G}\,\mathrm{d}x + \int_E \left(\frac{\partial}{\partial\theta_i}\log\lambda_Y\right)\left(\frac{\partial}{\partial\theta_j}\log\lambda_Y\right) g\bar{F}\,\mathrm{d}x,$$

for any $i,j = 1,\ldots,m$, where $I^{Z,\delta}(\theta_i,\theta_j)$ is the (i,j)th element of the Fisher information matrix for θ.

To obtain corresponding results for left censored data, $Z = \max(X,Y)$ and $\delta = I(X > Y)$, we use $\mu_X = f/F$, the reversed hazard rate of X. The analogy of (2) becomes

$$I^X(\theta) = \int_E \left(\frac{\partial}{\partial\theta}\log f\right)^2 f\,\mathrm{d}x = \int_E \left(\frac{\partial}{\partial\theta}\log\mu_X\right)^2 f\,\mathrm{d}x,$$

and Theorem 1 becomes

$$I^{Z,\delta}(\theta) = I^X(\theta) - \int_E \left(\frac{\partial}{\partial\theta}\log\mu_X\right)^2 f\,\bar{G}\,\mathrm{d}x + I^Y(\theta) - \int_E \left(\frac{\partial}{\partial\theta}\log\mu_Y\right)^2 g\,\bar{F}\,\mathrm{d}x.$$

Acknowledgements

This research was partly supported by an NSF grant while the first author was a graduate student in statistics at George Washington University. The work of the second author was completed when visiting Biostatistics Branch, Division of Cancer, Epidemiology and Genetics, NCI. The authors thank a referee for helpful suggestions that improve the manuscript.

References

Abdushukurov, I.A., Kim, L.V., 1987. Lower Cramér-Rao and Bhattacharya bounds for randomly censored observations. J. Soviet Math. 2171–2185.

Balakrishnan, N., 1985. Order statistics from the half logistic distribution. J. Statist. Comput. Simul. 20, 287–309.

Chen, Y.Y., Hollander, M., Langberg, N.A., 1982. Small-sample results for the Kaplan–Meier estimator. J. Am. Statist. Assoc. 77, 141–144.

Cheng, P.E., Lin, G.D., 1987. Maximum likelihood estimation of a survival function under the Koziol–Green proportional hazards model. Statist. Probab. Lett. 5, 75–80.

Csörgő, S., Horváth, L., 1981. On the Koziol–Green model for random censorship. Biometrika 68, 391–401.

Efron, B., Johnstone, I., 1990. Fisher information in terms of the hazard rate. Ann. Statist. 18, 38–62.

Hollander, M., Proschan, F., Sconing, J., 1985. Efficiency loss with the Kaplan–Meier estimator Technical report M707, Department of Statistics, Florida State University.

Koziol, J.A., Green, S.B., 1976. A Cramér-von Mises statistic for randomly censored data. Biometrika 63, 465–474.
Miller, R.G., 1983. What price Kaplan–Meier? Biometrics 39, 1077–1081.
Pawlitschko, J., 1999. A comparison of survival function estimators in the Koziol–Green model. Statistics 32, 277–291.
Prakasa Rao, B.L.S., 1995. Remarks on Cramér-Rao type integral inequalities for randomly censored data. In: Koul, H.L., Deshpande, J.V., (Eds.), Analysis of Censored Data, IMS Lecture Notes - Monograph Series, Vol. 27, Hayward, CA.
Stute, W., 1992. Strong consistency under the Koziol–Green model. Statist. Probab. Lett. 14, 313–320.
Zheng, G., 2001. A characterization of the factorization of hazard function by the Fisher information under type II censoring with application to the Weibull family. Statist. Probab. Lett. 53, 249–253.